

DAMIAN HERDA
Jagiellonian University
damian.t.herda@gmail.com

ON THE EFFECT OF PLURALIZATION ON THE NUMERALIZATION OF NOUNS IN ENGLISH AND POLISH: A CONTRASTIVE CORPUS-BASED STUDY

On the basis of corpus-derived data, the present paper examines the collocational patterns of the singular and the plural forms of a pair of etymologically and semantically related quantifying nouns (QNs), namely English *heap* and its Polish equivalent *kupa* ‘heap’. The primary aim is to determine their respective levels of numeralization, operationalized as the frequency of co-occurrence with animate and abstract N2-collocates in purely quantificational uses, in an attempt to establish whether, and to what extent, the addition of the plurality morpheme bears on the grammaticalization of a nominal of this kind into an indefinite quantifier. Following the observations arrived at by Brems (2003, 2011), the hypothesis is that pluralization should yield a facilitating effect on the numeralization of nouns referring to large quantities by amplifying their inherent scalar implications. The results demonstrate that whereas *heaps* indeed exhibits a higher percentage of such numeralized uses than *heap*, *kupy* ‘heaps’ has turned out to be grammaticalized in the quantifying function to a markedly lesser degree than *kupa* ‘heap’. It is argued that this apparently aberrant behaviour of *kupy* ‘heaps’ can nonetheless be elucidated in terms of the specificity of numeralization in Polish, since at its advanced, morphosyntactic stage, the process in question affects solely the singular (accusative) forms of QNs.

Keywords: *grammaticalization, numeralization, pluralization, quantifying nouns, English, Polish*

1. Introduction

Among the items which serve to convey imprecise quantitative assessments in English and Polish there is a vast group of so-called quantifying nouns (QNs), which have come to perform the function of indefinite quantifiers in a linguistic process known in the Polish literature as numeralization (cf. Schabowska 1962):

- (1) a. They had a {lot, load, heap...} of free time.
 b. Zarobili {kupe/mase/fure...} pieniędzy.
 ‘They earned a heap/mass/load of money.’

As noted by Brems (2003, 2011), it is possible for the singular forms of QNs pertaining to large numbers or amounts to have undergone different degrees of numeralization than the plural ones. Typically, it is the latter that are grammaticalized in the quantifying function to a larger extent, which leads her to the conclusion that the addition of the plurality morpheme amplifies the size implications inherent in the lexical make-up of such a QN, thus gratifying the speakers’ need for hyperbolic expressivity (Brems 2011: 155).¹ Based on extensive corpus-derived material, this paper investigates the effect of pluralization on the numeralization of a pair of etymologically and semantically related nouns in the two above-listed languages, namely English *heap* and Polish *kupa* ‘heap’. In view of the aforementioned observations arrived at by Brems (2003, 2011), the hypothesis advanced here is that *heaps* and *kupy* ‘heaps’ should exhibit a higher percentage of numeralized uses than *heap* and *kupa* ‘heap’, respectively.

The paper is organized as follows. Section 2 provides general information on the semantic and distributional characteristics of quantifiers. Section 3 sets the scene for the examination of the naturally-occurring data by providing an outline of the numeralization process in the light of grammaticalization theory. Section 4 introduces the research assumption, a description of the corpora from which the material has been extracted, an account of the applied methodology, as well as offers an analysis of the empirical data. The conclusions reached in the investigation are discussed in Section 5.

¹ Similarly, with reference to the initial stage of the grammaticalization of the English noun *lot* into a quantifier, Traugott and Trousdale (2013: 24) observe that the plural form *lots* occurred in contexts where the implicature of quantity was brought to the fore more frequently than the singular one, and it was likewise the former that extended its collocational scope to abstract nominals earlier than the latter.

2. Quantifiers: generalities

Quantifiers may be defined as linguistic exponents of quantitative evaluation (cf., among others, Laskowski 1984; Langacker 1991; Brems 2003, 2011; Radden and Dirven 2007; Delbecque and Verveckken 2014). Looked at from the point of view of semantics, quantifiers can be grouped into two primary types, namely *relative* and *absolute* quantifiers.²

Relative quantifiers “specify a quantity in relation to a reference mass” comprising the maximal instantiation of the pertinent category (Langacker 1991: 82), e.g. *all* and *most*. Therefore, Radden and Dirven (2007: 115) divide this category into full-set quantifiers, referring to each element of the set, viz. *all*, *every*, *each* and *any*, and subset quantifiers, focusing on individual elements of the set, viz. *most*, *half*, *some*. Full-set quantifiers, in turn, may be further subdivided into distributive quantifiers, viz. *each* and *every*, the collective quantifier *all* and the selective quantifier *any*. As regards absolute quantifiers, on the other hand, quantitative assessments are based “a scale with some implicit norm or standard” (Radden and Dirven 2007: 117), e.g. *many*, *much*, *(a) few*, *(a) little*, and *three*.

Following Langacker (1991: 84), it is likewise possible to classify quantifiers relative to the countability status of the nominals with which they co-occur. Accordingly, quantifiers may be categorized as follows: (i) the quantifier *one*, compatible solely with singular nouns, (ii) quantifiers used in relation to non-plural mass nouns, e.g. *much* and *(a) little*, and (iii) quantifiers appearing alongside plural nouns, e.g. *many* and *(a) few*.³ Also Radden and Dirven (2007: 131) distinguish between number quantification, which pertains to count nouns, and amount quantification, peculiar to mass nominals.

Apart from the above-discussed types, Quirk et al. (1985: 264) recognize the existence of an open-ended class of quantifying expressions, namely *phrasal quantifiers*, such as *lot(s) of*, *plenty of*, and *a great deal of*.⁴ Some of these can only occur with either count nouns, e.g. *a number of*, or mass nouns, e.g. *a large amount of*, whereas some are compatible with both types, e.g. *lot(s) of*. Analogously, Schabowska (1962), Langacker (1991: 88), and Radden and Dirven (2007: 134) note that in both English and Polish, there exists a large group of lexical elements capable of performing the function of vague quantification, such as *heap* in *heaps of time* and *kupa* ‘heap’ in *kupa roboty* ‘a heap of work’.

² There is, however, a slight terminological divergence to be noticed. In Radden and Dirven (2007), *relative* quantifiers are labelled as *set* quantifiers, whereas *absolute* quantifiers are referred to as *scalar* quantifiers.

³ Notably, Polish vague quantifiers are not as sensitive towards the countability of the nouns which they accompany as their English counterparts. For instance, both *malo* ‘few; a few’ and *duzo* ‘many; much’ are compatible with countable as well as uncountable nominals.

⁴ It is worth emphasizing here that in some of these quantifying phrases, the employment of an adjective is mandatory. Thus, it is only correct to say, for example, *a great/good deal of money*, whereas **a deal of money* (in the sense ‘a lot of money’) would be deemed ungrammatical in Standard English (cf. Quirk et al. 1985: 264).

Such items, known as quantifying nouns (cf. Biber et al. 1999: 252-255),⁵ acquire a purely quantificational sense as well as (some of) the distributional properties of canonical quantifiers in the diachronic process of numeralization.

3. Numeralization

As stated above, *numeralization* (Polish *numeralizacja*) is a term employed in the Polish literature with reference to the development of nouns into quantifiers (cf., among others, Schabowska 1962, 1967). Numeralization itself, in turn, may be thought of as an example of a larger linguistic phenomenon, namely *grammaticalization*, traditionally defined as a process whereby lexical items and constructions acquire grammatical, procedural meanings, and grammatical elements take on novel grammatical functions (Hopper and Traugott 2003: 1-2).

The initial phase of numeralization consists in a semantic generalization of the pertinent items (cf. Schabowska 1962; Hopper and Traugott 2003; Traugott and Trousdale 2013), so that the purely scalar inferences invited by nominals such as *heap* or *kupa* ‘heap’ become foregrounded to the detriment of other lexical features (Brems 2011: 231). As far as the two above-mentioned nouns are concerned, the backgrounded meaning component is that of a vertical arrangement of spatially contiguous entities:

- (2) *a heap of books* ‘a vertically arranged set of books’ > *a heap of books* ‘many books, regardless of their spatial distribution’

Langacker (1991: 88) observes that such grammaticalizing words have “taken on a different sense in which size becomes the most salient specification,” adding that “the notion of a discrete physical object has faded, leaving behind the conception of a schematically characterized mass.” This semantic change is also connected with subjectification, understood here as “a shift from the [QN] contributing to propositional content to expressing meaning that indexes speaker-relatedness, in that quantifier meaning involves a speaker assessing size relative to a scale” (Brems 2011: 231).

Importantly, the emergence of a purely quantificational sense leads to a collocational broadening of the relevant nominals (cf., among others, Brems 2003, 2011; Traugott and Trousdale 2013; Delbecque and Verveckken 2014), a phenomenon referred to as host-class expansion in Himmelmann (2004). In other words, having undergone meaning generalization, the items start to combine with N2s which do not satisfy the original selectional restrictions exhibited by particular QNs, i.e., as regards *heap* and *kupa* ‘heap’, with animate and abstract nouns as well as inanimate concrete noun collocates whose referents are not stackable:

⁵ In the Polish literature, the equivalent term is *rzeczownik ilościowy* ‘quantitative noun’ (cf. Schabowska 1967).

- (3) a heap of {newspapers/clothes/rubbish, etc.} > a heap of {water/people/time...}

What should likewise be emphasized in the context of the distributional changes affecting QNs is that in their purely quantificational uses, they can only be premodified by quantification-reinforcing adjectives, such as English *whole* and Polish *cały* ‘whole’, e.g. *a whole heap of time* and *cała kupa roboty* ‘a whole heap of work’ (cf. Brems 2011: 201). Numeralized nouns are further incompatible with other quantifiers (cf. Keizer 2007: 136), hence the ungrammaticality of (4a,b):

- (4) a. **trzy kupy roboty* ‘three heaps of work’, **każda kupa zabawy* ‘every heap of fun’,
**wiele kup pytań* ‘many heaps of questions’
 b. **three heaps of time*, **every heap of patience*, **many heaps of doubts*

Another stage of numeralization in English manifests itself in fluctuating verb agreement, which may serve as a formal indicator of a semantic change, as in some cases, it enables the ascertainment of whether a given element has been employed in its basic sense (5a), i.e. here, in relation to a collection of spatially bounded objects or an aggregate of a substance arranged vertically, or rather as an indefinite quantifier (5b), i.e. solely with reference to a considerable number or amount of what the concomitant nominal stands for (cf., among others, Langacker 1991; Brems 2003, 2011; Delbecq and Verweken 2014):

- (5) a. There was.SG a heap.SG of photographs.PL on the floor.
 b. There were.PL a heap.SG of photographs.PL on the floor.

According to Langacker (1991: 89), “[a]greement reflects the shift in profile from a discrete physical object that contains or constitutes a mass [...] to the quantified mass itself.” Notably, however, this syntactic reflex only applies to binominal syntagms which occur in the subject position, and exclusively to those in which the first and the second nominal element differ in number (Brems 2011: 129).

What may be considered an analogous change in Polish is the use of a QN in the accusative case when it occupies the subject position, with the verb being in the third person singular neuter form (cf. Schabowska 1962, 1967, 1970):⁶

⁶ This syntactic pattern is typical of Polish higher numerals (for a discussion of the peculiarities of the syntax of Polish quantifiers, see, among others, Przepiórkowski 2004):

(i) Pięć osób przyszło/ *przyszły na przyjęcie.
 five person.PL.FEM.GEN come.PST.3.SG.NEUT come.PST.3.PL.FEM on party
 ‘Five people came to the party.’

- (6) Kupe ludzi tam było.
 heap.SG.FEM.ACC people.PL.MASC.GEN there be.PST.3.SG.NEUT
 ‘There were a heap of people there.’

The above-described morphosyntactic changes, together with the restricted premodification patterns of QNs, can be looked at as a sign of their (incomplete) decategorialization (Brems 2011: 232; cf. also Hopper 1991; Heine 2003), which means that nominal elements affected by numeralization gradually become devoid of the properties of the class of nouns, and instead acquire an increasing number of characteristics of the category of quantifier (Brems 2011: 111).⁷

In regards to semantics, numeralization at an advanced stage may lead to the pertinent items losing their original meaning, a phenomenon referred to in the grammaticalization framework as *semantic attrition/bleaching/reduction* (cf. Lehmann 1985; Sweetser 1988; Heine 2003; Hopper and Traugott 2003), as has been the case with, for instance, the Polish vague quantifier *trochę* ‘a little’ (cf. Schabowska 1970), etymologically related to Proto-Slavonic **troska* ‘bit; chip; scrap’ (SEJP, p. 642). Interestingly, Brems (2011: 203) notes an analogous tendency for the English QN *lot*, whose quantifier attestations (‘many/much’) significantly outnumber the partitive ones (‘a unit of’, as in *three lots of land*) in actual language use.

4. Analysis of empirical data

4.1. Hypothesis

In consonance with Brems (2003, 2011), it is assumed that the plural forms of the analysed QNs should display higher degrees of numeralization, operationalized here as the frequency of use in the quantifying function in relation to animate and abstract N2-collocates, than the singular ones, the reason being that “[g]rammatical number adds to the hyperbolic value” (Brems 2011: 155). Additionally, pluralized nominals resemble mass nouns in that both are conceptualized as unbounded (cf. Jackendoff 1991; Corbett 2004: 79-80), which likewise enhances the implicature of magnitude, already present in the singular form of a ‘large size’ QN.

Also crucial in this context is the denotational affinity between *heap* and *kupa* ‘heap’, since, as Brems (2011: 203) remarks, in the case of nouns encoding more specific, cohesive arrangements, such as *bunch*, pluralization may achieve a different effect, i.e. instead of simply pointing to an abundance of the referent(s) of the accompanying nominal, *bunches* is more likely to indicate

⁷ Following Brems (2003, 2011) and Delbecq and Verweyken (2014), QNs, at least at the early stages of their grammaticalization, can be preceded by definite determiners, as is also the case with some canonical quantifiers, e.g. *many in the many questions that were asked*.

a number of separate clusters composed thereof. At stake here is the fact that the concept of *heap* and *kupa* ‘heap’ “lacks [the] idea of constructional solidity” and “profiles an undifferentiated mass with a more chaotic internal consistency” (Brems 2011: 156), as suggested by the dictionary definitions cited below:⁸

heap

- (i) an untidy pile of something
- (ii) (*informal*) a lot of something
- (iii) (*informal, humorous*) a car that is old and in bad condition (OALD, p. 719)

kupa

- (i) many things lying, heaped, thrown together; stack, pile
- (ii) a great number of people or animals grouped together; crowd, mob, herd
- (iii) a great amount or number of something; a lot, plenty
- (iv) (*colloq.*) excrement (USJP, p. 565)

On the basis of the relevant etymological information provided in CEDEL (pp. 710-711), OnED, and SEJP (p. 274), it can be observed that *heap* and *kupa* ‘heap’ are moreover genetically related, and most probably derive from the Indo-European base **geu-p* ‘bend; arch; vault’. The former is a continuation of West Germanic **haupaz*, whereas the latter stems from Old Slavonic **kupŭ*.

4.2. Corpora

The Polish data were extracted from the 300 million word National Corpus of Polish (NKJP). The English material, on the other hand, comes from a number of smaller corpora, namely the 100 million word British National Corpus (BNC), the TIME Magazine Corpus of American English (TIME), equally large as the BNC, the 50 million word Strathy Corpus of Canadian English (CAN), and finally, the Corpus of Online Registers of English (CORE), also consisting of approximately 50 million tokens. Notably, the total size of all the above-listed English corpora corresponds to that of the Polish one, and the entire material is register-diversified. Besides, the English data represent different regional varieties, which allows us to obtain a more accurate, balanced description of the general use of *heap* and *heaps*.

4.3. Methodology

The first step of the data collection process involved the extraction of all the attestations of *heap* and *heaps* as well as *kupa* ‘heap’ and *kupy* ‘heaps’ in the binominal construction which can be schematically presented as ‘N1 of N2’

⁸ The translations of the Polish dictionary definitions as well as of the corpus examples are mine, and so is the bolding and underlining used in the original texts cited here – DH.

in English, and ‘N1 N2.GEN’ in Polish, wherein the N1-slot is occupied by the QN, and N2 stands for the concomitant nominal, together with its potential pre-modifiers, in relation to whose referent a quantitative assessment is being conveyed.⁹

Subsequently, the extracted material was edited manually, which involved the elimination of the examples in which the N1 had been used in a sense other than the arrangement/quantity one, including instances of what Lehrer (1986: 111) calls metaphorical comparison classifier uses, e.g. *a heap of a car*, as well as the removal of occasional doublets. Also excluded were the syntagms in which the N2-slot was occupied by a pronoun, since it was not always possible to determine its actual reference. The examples eliminated at this point were then not taken into account in the quantitative examination of the data.

The core of the analysis was the determination of the extents to which the items under scrutiny have numeralized by grouping their N2-collocates into the following categories: (i) concrete, (ii) animate, and (iii) abstract. Crucially, the last two types of N2s were then searched for pure quantifier uses (PQUs), i.e. instances in which the N1 may be felicitously replaced with a more canonical quantifier without distorting the intended meaning (cf. Brems 2011: 229), based on the assumption that the more numerous and semantically diverse the animate and, in particular, abstract noun collocates, the higher degree of numeralization. This required the filtering out of instances of, among other things, valuing quantifier uses (Brems 2011: 175),¹⁰ in which the N1 functions as a negatively charged collective noun rather than a schematic quantifier, e.g.:¹¹

- (7) Dla mnie to **kupa bredni**. [NKJP]
 ‘For me, it’s a heap of idiocies.’

The final step consisted in establishing the proportions of the PQUs of *heap*, *heaps*, *kupa* ‘heap’, and *kupy* ‘heaps’ involving animate and abstract N2s in relation to all of their nominal collocates, including the concrete ones. In what follows, the results of the quantitative investigation are presented and briefly discussed.

⁹ The Polish data were extracted by means of the Poliqarp search engine.

¹⁰ Such qualifying uses are peculiar to the singular forms of QNs such as *heap*, as they “have the advantage of singling out one set of entities in terms of a shared qualitative feature and of keeping their focus on that one bounded set which is then usually negatively valued” (Brems 2011: 203).

¹¹ In some contexts, however, it is problematic to distinguish between purely quantificational and evaluative uses, especially when the N1 can be equally felicitously replaced with a standard quantifier as well as a negatively loaded collective nominal. Such ambiguous instances were nonetheless classified as (potential) PQUs.

4.4. Results

The following table shows the results of the quantitative analysis of the degree to which each of the analysed forms has numeralized, as reflected in the proportion of its PQUs involving animate and abstract noun collocates in relation to the total number of its relevant attestations in the binominal 'N2 (of) N2(.GEN)''-construction:

Table 1. PQUs of *heap(s)* and *kupa/kupy* 'heap/heaps' with animate and abstract collocates

QN		Total no. of attestations	PQUs with animate and abstract N2s (N2-types)	% of PQUs with animate and abstract N2s
English	<i>heap</i>	692	110 (13 animate, 97 abstract)	15.90%
	<i>heaps</i>	433	106 (22 animate, 84 abstract)	24.48%
Polish	<i>kupa</i> 'heap'	1379	339 (156 animate, 183 abstract)	24.58%
	<i>kupy</i> 'heaps'	164	19 (19 animate, 0 abstract)	11.59%

As demonstrated above, whereas the case of English *heap(s)* indeed corroborates the hypothesis specified in section 4.1, the Polish data reveal that when confronted with cross-linguistic evidence, the research assumption does not hold. The subsequent parts of the paper offer a qualitative discussion of the results, illustrated with a number of representative corpus examples.¹²

4.4.1. *Heap(s)*

a) *Heap*

When employed in relation to animate nouns, *heap of* typically functions as a quantifier substitutable with *many* or *lot(s) of*. Moreover, in a majority of instances involving human N2-collocates, *heap*'s original lexical feature of spatio-temporal contiguity is considerably backgrounded:

- (8) "We are not meeting Julia's education targets, or words to that effect. We have **a heap of young people** who are not earning or learning". Etc. # The number of young people aged 18 to 24 working, studying or training fell 3.8 per cent to 72.5 per cent, the report showed. [CORE]

¹² Due to space limitations, it is here impossible to provide an exhaustive list of the N2-collocates of the analysed items identified in the data.

In (8), the young people referred to are most plausibly spread all over a country, and *heap* solely conveys the speaker's quantitative assessment rather than classifies the animates in terms of spatial distribution. Notably, in some cases, the quantifier reading is reinforced by verb concord with the N2, which may be regarded as a formal reflex of the grammaticalized status of *heap of*, comparable to that of *lot(s) of*:

- (9) There *were* **a heap of experience people** within the district. [CORE]
- (10) There *are* **a heap of manufacturers** now doing cross-over type vehicles, and they're definitely a popular choice. [CORE]

In the abstract domain, the quantifier *heap of* displays a remarkably wide collocational scope, as it habitually co-occurs with, among other semantic N2-types, nouns standing for psychological phenomena (11), activity-denoting nominals (12), as well as N2s referring to pragmatic acts (13):

- (11) It ought to be a very pleasant adventure even if it should require **a heap of patience**... [TIME]
- (12) Sure would save **a heap of mucking about**. [CORE]
- (13) Before any additional Alpha licencees are revealed, DEC says it has **a heap of important VAX announcements** to make. [BNC]

b) *Heaps*

When accompanied by animate N2s, *heaps* is predominantly employed as an indefinite quantifier, bleached of its lexical features, such as the requirement that the referents of the concomitant nominal be characterized by physical closeness, other than the purely scalar implications. The lack of spatial proximity of the N2-referents is especially evident in the general statement made in (14), since the kids in question are spread all over the world:

- (14) Most populations have **heaps of kids** because they will look after them when they're old. [CORE]

As is also the case with *heap*, a vast group of the abstract collocates of *heaps* comprises nominals denoting a wide range of psychological concepts. By and large, such N2s can be divided into two categories: those which display a positive colouring (15), and those which are negatively charged (16):

- (15) Now I have **heaps of confidence** as a writer and knowledge about publishing [...]. [CORE]
- (16) So how is an aspiring poet to deal with **such heaps of discouragement**? [CAN]

Notably, in (16), the quantity implications of *heaps* are additionally strengthened by the intensifier *such*. Semantically akin to the psychological nouns above are N2s such as the following ones, which can be broadly described as referring to various sensations, principally pleasant experiences:

- (17) It sounds (and looks) like they had **heaps of fun**, and even met the Prime Minister of New Zealand, John Key! [CORE]
- (18) FRIDAY: Cruise north again to the pretty mainland resort of Parga, set in a beautiful bay surrounded by mountains. A lively, bustling place with **heaps of atmosphere**. [BNC]
- (19) The Zind Humbrecht 1988 Pinot Blanc is quintessential Alsace: intense with an inviting spring-flower aroma and **heaps of fresh grapey-fruity flavour**. [BNC]

In addition, analogously to *heap*, *heaps* regularly co-occurs with event-denoting N2s, as illustrated by (20) and (21):

- (20) He signs endless autographs, does tons of interviews, shows up at **heaps of charity events**, never trash talks on the court and tidies up the day before the cleaning service comes. [TIME]
- (21) Oh well! I've done **heaps of fashion shows** in my life (so far) so walking the runway for charity is a total pleasure. [CORE]

Moreover, *heaps* has been observed to quite frequently quantify over nouns which denote the concepts of *time* (22) and *space/room* (23-24):

- (22) There's **heaps of time**. [BNC]
- (23) They're outdoors, have **heaps of space** and a natural environment. [CORE]
- (24) Oh a million times yes. YES. As you say, there's **heaps of room**. [CORE]

What the above discussion demonstrates is that in their PQUs, *heap* and *heaps* tend to co-occur with a wide range of semantically varied nominals, a finding which testifies to the advanced semantic schematicity of both forms, mirrored in their collocational freedom comparable to that of the standard English quantifiers *many*, *much*, and *lot(s) of*.

4.4.2. *Kupa/Kupy* 'heap/heaps'

a) *Kupa* 'heap'

Similarly to *heap* and *heaps*, *kupa* 'heap' is capable of quantifying over nominals whose referents are not marked by spatio-temporal closeness, with the arrangement feature intrinsic to its lexical profile being backgrounded:

- (25) Obcy na koniach rzucają się w oczy, a Słowik ma w mieście **kupę informatorów**. [NKJP]
 ‘Strangers riding on horses are quite conspicuous, and Słowik has a heap of informants in town.’
- (26) Coś nie tak? Przecież po mnie **kupa ludzi** to przeglądała. Ktoś kwestionuje opinię? [NKJP]
 ‘Something’s wrong? A heap of people have looked through it after me, mind you. Is anyone questioning the opinion?’

In (25), the context implies that the informants referred to are spread over the whole area of a city, hence the assumed lack of their physical proximity. Analogously, in (26), the N2-referents are most plausibly not characterized by either spatial or temporal contiguity, i.e. the document at issue has been inspected by many people over a certain time-span, perhaps by one person at a time. In a number of cases, the grammaticalized status of *kupa* ‘heap’ is reflected in the verb in the past tense being in the neuter, not feminine, with the very item in its accusative form occupying the subject position, as shown in (27-28):

- (27) Minęło już **kupę lat** od czasu, kiedy zarobiłem ostatniego dolara dla Acorn. [NKJP]
 ‘A heap of years have passed since I earned the last dollar for Acorn.’
- (28) no i tam właśnie z tego tego biura konsaltingu **kupę ludzi** wiesz się porozchodziło [NKJP]
 ‘And, you know, a heap of people have left this consulting office.’

As for its abstract N2-collocates, *kupa* ‘heap’ has been found to co-occur with a wide range of nominals representative of diverse semantic domains. It is nonetheless possible to discern certain collocational preferences. For instance, the QN under analysis habitually collocates with the N2 *czas* ‘time’ (29) as well as the activity-denoting nominals *robota* ‘work’ (30) and *śmiech* ‘laughter’ (31):

- (29) Już i tak zmitrężyłem **kupę czasu**. [NKJP]
 ‘I’ve already wasted a heap of time, anyway.’
- (30) Mimo że z dziećmi jest **kupa roboty**, to i tak kondycję człowiek traci. [NKJP]
 ‘Although taking care of children entails a heap of work, one loses good form, anyway.’
- (31) Ten konkurs to była superzabawa na scenie, **kupa śmiechu** i miłe, bardzo miłe, najmiłsze zaskoczenie! [NKJP]
 ‘That competition was super fun on stage, a heap of laughter and a nice, very nice, the nicest surprise!’

Moreover, as is also the case with *heap*, *kupa* ‘heap’ frequently co-occurs with psychological nouns, which can be broadly grouped into those negatively loaded (32) and positively coloured (33):

- (32) Mieliśmy z załatwieniem tych butów **kupe kłopotu**. [NKJP]
 ‘We had a heap of trouble fixing these shoes.’
- (33) Pieniądze dają **kupe szczęścia**, byle wiedzieć, co z nimi robić. [NKJP]
 ‘Money can give you a heap of happiness, provided that you know how to use it.’

b) *Kupy* ‘heaps’

As can be seen in Table 1, all of the PQUs of *kupy* ‘heaps’ involve animate N2-collocates. Nevertheless, it is typically impossible, due to a scarcity of contextual clues, to determine whether the form is meant to indicate just a multitude of the N2-referents, or whether it points to a number of *groups* composed thereof. Sometimes, it is the quantification-reinforcing adjective *cały* ‘whole’ that suggests the quantifier reading:¹³

- (34) Ale pomimo nabożeństw ulice roiły się od ludzi, zaś przy kramach pod Bramą Krakowską, z których roznosiły się swędy prażonych kiełbas, zbierały się **całe kupy pospólstwa**. [NKJP]
 ‘Yet, despite the church services, the streets were swarming with people, and whole heaps of populace were gathering at the stalls near the Kraków Gate, whence the smell of roasted sausages was wafting.’

Even in such cases, however, we may observe a high level of persistence of the QN’s original lexical feature of spatio-temporal contiguity of the N2-referents: in (34), for instance, it is clear that the pertinent individuals occupied the same area within the same temporal frame.

Yet, as against *kupa* ‘heap’, which has been shown to co-appear with a variety of abstract nouns, *kupy* ‘heaps’ has not developed quantifier uses with this kind of nominals at all. The sole example involving an abstract N2-collocate detected in the data is in fact just the plural version of the semi-lexicalized phrase *kupa nieszczęścia* ‘poor thing; lit.: a heap of misfortune’, employed metaphorically in relation to an individual deemed miserable by the speaker:

- (35) Jeśli wszyscy nie podskoczyliśmy z wrażenia, to tylko dlatego, że byliśmy trójką bezsilnych **kup nieszczęścia**. [NKJP]
 ‘If we all did not jump with fright, it was only because we were three helpless heaps of misfortune [=poor things].’

¹³ Curiously, *heaps*, as against *heap*, is not modified by the adjective *whole* at all in the scrutinized corpus data, which suggests that in English, pluralization of ‘large size’ QNs typically suffices to increase their quantifier potential, so that they no longer need to rely on such adjectival intensifiers.

5. Conclusion

The results obtained in this study demonstrate that cross-linguistically, the hypothesis which postulates a facilitating effect of pluralization on the grammaticalization of nouns such as *heap* into indefinite quantifiers cannot be sustained. Whereas the English data indeed corroborate the assumption, since it is *heaps* that has numeralized to a higher degree than *heap*, the case of Polish *kupa* ‘heap’ indicates that pluralization does not invariably enhance the scalar implications of a QN.¹⁴ Moreover, whereas both *heap*, *heaps* and *kupa* ‘heap’ exhibit a semantically varied collocational scope, including animate as well as numerous count and non-count abstract noun collocates, *kupy* ‘heaps’ has been shown not to quantify over abstract nominals whatsoever, and even some of the instances in which the latter co-occurs with animate N2s are in fact ambiguous between the pure quantifier (‘many; lot(s) of’) and the partitive (‘groups of’) reading.

A possible explanation for the above findings may be that in Polish, only the singular forms of QNs, whether pointing to small or large quantities, in the accusative case are recruited in the process of numeralization at its morphosyntactic stage (cf. examples (27-28)).¹⁵ In other words, the accusative form of a QN eventually becomes specialized as a quantifier, which is best exemplified by the item *trochę* ‘a little’, which comes from the now non-existent feminine noun *trocha* ‘small amount’, genetically connected with Proto-Slavonic **troska* ‘bit; chip; scrap’ (SEJP, p. 642), yet has lost all of its nominal properties, such as the capability of being pluralized, inflected for case, and modified by adjectives. Interestingly, a number of Polish QNs expressing large numbers or amounts, e.g. *masa* ‘masa’, *fura* ‘load’, as well as the scrutinized item *kupa* ‘heap’, are also feminine nouns taking the *-ę*-ending in their singular accusative forms, which suggests that the development of nominals into vague quantifiers in Polish may be fuelled by analogy to the above-mentioned element *trocha*

¹⁴ A preliminary corpus examination of the collocability of the singular and plural forms of the Polish QNs *fura* ‘load’ and *masa* ‘mass’ reveals the same tendency: it is *fura* ‘load’ and *masa* ‘mass’ that function quantificationally in relation to semantically varied N2-collocates substantially more frequently than *fury* ‘loads’ and *masy* ‘masses’, respectively, an observation which strengthens the conclusions arrived at in the present study.

¹⁵ Likewise, only the singular accusative forms of Polish QNs may be affected by adverbialization, which constitutes another stage in the grammatical evolution of such items:

- (i) Ostatnio masę.SG.FEM.ACC podróżował.
‘He has travelled a mass [=a lot] recently.’
- (ii) *Ostatnio masy.PL.FEM.ACC podróżował.
‘He has travelled masses [=a lot] recently.’

In English, by contrast, both singular and plural forms of QNs have been observed to be capable of functioning as degree adverbs, e.g. *a load more* ‘much more’, *heaps better* ‘much better’ (cf. De Clerck and Brems 2016).

‘small amount’, which was entrenched in the quantifying function already in Old Polish (cf. SSP, p. 186).

Another fact which may be seen as underlying the lack of systematic numeralization of *kupy* ‘heaps’ as well as the plural forms of other Polish QNs (cf. footnote 14) is that in Polish, plural morphology performs the function of unitization to a greater extent than is the case in English. By way of illustration, while it is customary for speakers of Polish to use forms such as *drewna* ‘woods (pieces of wood)’, *sery* ‘cheeses (portions of cheese)’, or *soki* ‘juices (portions of juice)’, analogous uses in English would be deemed incorrect by most speakers of the latter language (cf. Willim 2006: 59). That pluralization of English ‘large size’ QNs, as opposed to that of Polish ones, generally yields a massifying effect, thus facilitating numeralization, can likewise be linked with potential differences in the extent of application of the so-called intensificative plural in the two languages (cf. Corbett 2004: 238), exemplified by *sands* (as in *the sands of the desert*) and *wody* ‘waters’ (as in *wody oceanu* ‘the waters of the ocean’), both of which stand not for portions or kinds of the pertinent substance, but for *large amounts* thereof, so that the plural morpheme plays the semantic role of an indefinite quantifier expressing an abundance of what the associated nominal refers to (Willim 2007: 184). Given the results of the empirical analysis reported here, this special employment of plural morphology may be assumed to exhibit a higher frequency of occurrence in English, yet this hypothesis requires further investigation.

References

- Biber, D., S. Johansson, G. Leech, S. Conrad and E. Finegan 1999. *Longman grammar of spoken and written English*. London: Longman.
- Brems, L. 2003. Measure noun constructions: an instance of semantically-driven grammaticalization. *International Journal of Corpus Linguistics* 8(2): 283-312.
- Brems, L. 2011. *The layering of size noun and type noun constructions in English*. Berlin/Boston: De Gruyter Mouton.
- Corbett, G.G. 2004. *Number*. Cambridge: Cambridge University Press.
- De Clerck, B., and L. Brems 2016. Size nouns matter: a closer look at *mass(es) of* and extended uses of SNs. *Language Sciences* 53: 160-176.
- Delbecque, N., and K.D. Verveckken 2014. Conceptually-driven analogy in the grammaticalization of Spanish binominal quantifiers. *Linguistics* 52(3): 637-684.
- Heine, B. 2003. Grammaticalization. In R. Janda and B. Joseph (eds.), *The handbook of historical linguistics*, 575-601. Oxford: Blackwell.
- Himmelmann, N.P. 2004. Lexicalization and grammaticization: Opposite or orthogonal? In W. Bisang, N.P. Himmelmann and B. Wiemer (eds.), *What makes grammaticalization? A look from its fringes and its components*, 21-42. Berlin/New York: Mouton de Gruyter.

- Hopper, P.J. 1991. On some principles of grammaticization. In E.C. Traugott and B. Heine (eds.), *Approaches to grammaticalization: Volume 1*, 17-35. Amsterdam/Philadelphia: John Benjamins.
- Hopper, P.J., and E.C. Traugott 2003. *Grammaticalization*. Second Edition. New York: Cambridge University Press.
- Jackendoff, R. 1991. Parts and boundaries. *Cognition* 41: 9-45.
- Keizer, E. 2007. *The English noun phrase: the nature of linguistic categorization*. Cambridge: Cambridge University Press.
- Langacker, R. 1991. *Foundations of cognitive grammar: Volume 2: descriptive application*. Stanford: Stanford University Press.
- Laskowski, R. 1984. Liczebnik. In R. Grzegorzczkowska, R. Laskowski and H. Wróbel (eds.), *Gramatyka współczesnego języka polskiego: Morfologia*, 283-293. Warszawa: Państwowe Wydawnictwo Naukowe.
- Lehmann, C. 1985. Grammaticalization: synchronic variation and diachronic change. *Lingua e Stile* 20: 303-318.
- Lehrer, A. 1986. English classifier constructions. *Lingua* 68: 109-148.
- Przepiórkowski, A. 2004. O wartości przypadku podmiotów liczebnikowych. *Biuletyn Polskiego Towarzystwa Językoznawczego* 60: 133-143.
- Quirk, R., S. Greenbaum, G. Leech and J. Svartvik 1985. *A comprehensive grammar of the English language*. London/New York: Longman.
- Radden, G., and R. Dirven 2007. *Cognitive English grammar*. Amsterdam: John Benjamins.
- Schabowska, M. 1962. O formalnej numeralizacji rzeczowników. *Język Polski* 42(2): 116-124.
- Schabowska, M. 1967. *Rzeczowniki ilościowe w języku polskim*. Wrocław: Ossolineum.
- Schabowska, M. 1970. Liczebnik nieokreślony *trochę*. *Język Polski* 50(2): 110-118.
- Sweetser, E. 1988. Grammaticalization and semantic bleaching. In S. Axmaker, A. Jaisser and H. Singmaster (eds.), *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*, 389-405. Berkeley: Berkeley Linguistics Society.
- Traugott, E.C., and G. Trousdale 2013. *Constructionalization and constructional changes*. Oxford: Oxford University Press.
- Willim, E. 2006. *Event, individuation and countability: a study with special reference to English and Polish*. Kraków: Jagiellonian University Press.
- Willim, E. 2007. Liczba jako kategoria gramatyczna i pojęciowa. *Studia Linguistica Universitatis Iagellonicae Cracoviensis* 124: 175-192.

Corpora

- BNC = Davis, M. (n.d.). British National Corpus (BYU-BNC) [<http://corpus.byu.edu/bnc/>; accessed October 5, 2017].
- CAN = Davis, M. (n.d.). Corpus of Canadian English (Strathy) [<http://corpus.byu.edu/can/>; accessed October 5, 2017].

CORE = Davis, M. (n.d.). CORE Corpus of Online Registers of English [<http://corpus.byu.edu/core/>; accessed October 5, 2017].

NKJP = Bańko, M., R. L. Górski, B. Lewandowska-Tomaszczyk, M. Łaziński, P. Pęzik and A. Przepiórkowski. (2008-2012), Narodowy Korpus Języka Polskiego [<http://nkjp.pl/poliqarp/>; accessed October 5, 2017].

TIME = Davis, M. (n.d.). TIME Magazine Corpus of American English [<http://corpus.byu.edu/time/>; accessed October 5, 2017].

Dictionaries

CEDEL = Klein, E. 1966. *A comprehensive etymological dictionary of the English language: Volume I: A–K*. Amsterdam/London/New York: Elsevier.

OALD = Turnbull, J. (ed.) 2010. *Oxford advanced learner's dictionary of current English: Eighth Edition*. Oxford: Oxford University Press.

OnED = Harper, D. 2001-2018. *Online etymology dictionary* [<http://www.etymonline.com/>; accessed September 30, 2017].

SEJP = Boryś, W. 2005. *Słownik etymologiczny języka polskiego*. Kraków: Wydawnictwo Literackie.

SSP = Urbańczyk, S. (ed.) 1953-2002. *Słownik staropolski: T. IX: Ściadły–Używowanie*. Kraków: Polska Akademia Nauk.

USJP = Dubisz, S. (ed.) 2003. *Uniwersalny słownik języka polskiego: Tom 2: H–N*. Warszawa: Wydawnictwo Naukowe PWN.